



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
-----------------	-------------	----------------------	---------------------	------------------

10/785,199

02/25/2004

Misty Azara

CQ10218

3364

23493 7590 05/13/2009  
SUGHRUE MION, PLLC  
2100 Pennsylvania Avenue, N.W.  
Washington, DC 20037

EXAMINER

COLUCCI, MICHAEL C

ART UNIT

PAPER NUMBER

2626

NOTIFICATION DATE

DELIVERY MODE

05/13/2009

ELECTRONIC

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

USPTO@sughrue.com  
USPatDocketing@sughrue.com

<b>Office Action Summary</b>	<b>Application No.</b> 10/785,199	<b>Applicant(s)</b> AZARA ET AL.	
	<b>Examiner</b> MICHAEL C. COLUCCI	<b>Art Unit</b> 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☒ Responsive to communication(s) filed on 02/04/2009.
- 2a) ☒ This action is **FINAL**.                      2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 1-12, 15-26 and 30 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-12, 15-26 and 30 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \*    c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)          | 4) <input type="checkbox"/> Interview Summary (PTO-413)           |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____                                      |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)          | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date _____  | 6) <input type="checkbox"/> Other: _____                          |

## **DETAILED ACTION**

### ***Response to Arguments***

1. Applicant's arguments filed 02/04/2009 have been fully considered but they are not persuasive.

#### **Argument (page 10):**

#### **Response to argument & corresponding amendment to claims:**

When considered in view of the specification of the present invention, Examiner believes that the combined teaching of Marcu in view of Lee appears to teach the limitation of claims 1, 15, and 30 as amended. Further, the present invention describes that "the theory of discourse analysis may include any theory of discourse analysis capable of identifying discourse functions in a text" (present invention spec. [0021] & Fig. 2, first determining a discourse theory, or any theory). Marcu demonstrates this by using a rhetorical structure theory (RST). Furthermore, this RST is consistent with the present invention, wherein the present invention teaches "marking discourse level structures such as the nucleus and satellite distinction described in Rhetorical Structures Theory" (present invention spec. [0021]). Marcu clearly demonstrates an RST theory for discourse analysis (Marcu [0021]). Marcu goes on to describe nuclei and satellite concepts in relation to the RST, wherein the text is thoroughly understood ([0077-0078]).

Art Unit: 2626

Additionally, given the teaching of Marcu relative to an RST, Lee further describes user intervention, wherein a user can change gender, age, and speech rate of the synthesized speech (Lee Col. 7 lines 10-17 & Fig. 1).

Further, Lee teaches the ability to control and adjust prosodic parameters for speech synthesis (Previously cited in the last office action Col. 2 lines 29-49 & Fig. 1).

Though, Lee deals with pictures, the multimedia/images/pictures analysis is separate from the act of receiving input text, identifying phonemic data, and adjusting features for synthesis (i.e. TTS). This concept is well known, and thus Lee merely uses a well known concept to improve multimedia synchronization. The well known teaching of Lee with respect to receiving input text (word, sentence, pattern, etc.), identifying phonemic data, and adjusting features for synthesis allows for an improvement to the combined teachings of Shriberg and Marcu, wherein a system that analyzes text based on a model of discourse analysis can now synthesize speech based on adjusted parameters. This is merely an act of text to speech synthesis based on adjusted feature values.

***Claim Rejections - 35 USC § 103***

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

3. Claims 1-12, 15-26, and 30 are rejected under 35 U.S.C. 103(a) as being unpatentable over “Can Prosody Aid the Automatic Classification of Dialog Acts in Conversational Speech?” (hereinafter Shriberg) in view of Marcu et al. US 20020046018 A1 (hereinafter Marcu) and further in view of Lee et al. US 6088673 A (hereinafter Lee).

Re claims 1, 15, and 30, Shriberg teaches a method of synthesizing speech (Page 5) using discourse function level prosodic features (Pages 14-18) comprising the steps of:

determining discourse functions in the input text the discourse functions being determined based on a mapping between basic discourse constituents of the determined theory of discourse analysis and a plurality of discourse functions (Pages 8-13);

determining a model of discourse function level prosodic features (Pages 14-18);

However, Shriberg fails to teach determining a theory of discourse analysis from plurality of theories of discourse analysis;

Marcu teaches a channel-based summarization process 1700 is to receive the input text. Although the embodiment described above uses sentences as the input text, any other text segment could be used instead, for example, clauses, paragraphs, or entire treatises. Next, in step 1704, the input text is parsed to produce a syntactic tree in the style of FIG. 11, which is used in step 1706 as the basis of generating multiple possible solutions (e.g., the shared-forest structure described above). If a whole text is given as input, the text can be parsed to produce a discourse tree, and the algorithm described here will operate on the discourse tree (Marcu [0220-0221]).

Further, Marcu teaches a discourse structure for an input text segment (e.g., a clause, a sentence, a paragraph or a treatise) is determined by generating a set of one or more discourse parsing decision rules based on a training set, and determining a discourse structure for the input text segment by applying the generated set of discourse parsing decision rules to the input text segment (Marcu [0010]).

Furthermore, Marcu teaches generating the set of discourse parsing decision rules may include iteratively performing one or more operations (e.g., a shift operation and one or more different types of reduce operations) on a set of edus to incrementally build the annotated text segment associated with the set of edus. The different types of reduce operations may include one or more of the following six operations: reduce-ns, reduce-sn, reduce-nn, reduce-below-ns, reduce-below-sn, reduce-below-nn. The six reduce operations and the shift operation may be sufficient to derive the discourse tree of any input text segment (Marcu [0012]).

Marcu clearly demonstrates an RST theory for discourse analysis (Marcu [0021]). Marcu goes on to describe nuclei and satellite concepts in relation to the RST, wherein the text is thoroughly understood ([0077-0078]).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Shriberg to incorporate determining a theory of discourse analysis from plurality of theories of discourse analysis as taught by Marcu to allow for the proper rules to analyze input text, wherein the type of input (phrases, sentences, words, etc.) determine how to determine the structure of text such as rhetorical analysis (Marcu [0012]).

However, Shriberg in view of Marcu fail to teach  
determining input text;  
determining adjusted synthesized speech output based on the discourse functions, the model of discourse function level prosodic features (pages 14-18), and the input text

Lee teaches a TTS for interlocking with multimedia according to the present invention comprises a multimedia information input unit for organizing text, prosody, the information on synchronization with moving picture, lip-shape, and the information such as individual property; a data distributor by each media for distributing the information of the multimedia information input unit into the information by each media; a language processor for converting the text distributed by the data distributor by each media into phoneme stream, presuming prosody information and symbolizing the information; a

Art Unit: 2626

prosody processor for calculating a value of prosody control parameter from the symbolized prosody information using a rule and a table; a synchronization adjuster for adjusting the duration of the phoneme using the synchronization information distributed by the data distributor by each media; a signal processor for producing a synthesized speech using the prosody control parameter and data in a synthesis unit database; and a picture output apparatus for outputting the picture information distributed by the data distributor by each media onto a screen (Lee Col. 2 lines 29-49 & Fig. 1).

Lee further describes user intervention, wherein a user can change gender, age, and speech rate of the synthesized speech (Lee Col. 7 lines 10-17 & Fig. 1).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Shriberg in view of Marcu to incorporate determining input text and determining adjusted synthesized speech output based on the discourse functions, the model of discourse function level prosodic features, and the input text as taught by Lee to allow for the proper rules to analyze input text, wherein prosody control is established such as phonemic features of text in order to modify output speech synthesis based on input text, and words, sentences, or textual patterns in order to adapt in a changing environment (Lee Col. 7 lines 10-17 & Fig. 1).

Re claims 2 and 16, Shriberg teaches the method of claim 1, wherein the discourse functions are determined based on the determined theory of discourse analysis (Pages 8-13).



Re claims 3 and 17, Shriberg fails to teach the method of claim 2, in which the theory of discourse analysis is at least one of: the Linguistic Discourse Model, the Unified Linguistic Discourse Model, Rhetorical Structures Theory, Discourse Structure Theory and Structured Discourse Representation Theory;

Re claims 4 and 18, Shriberg teaches the method of claim 1, wherein the output information (Pages 4-5, Why Use Prosody?) is at least one of text information and application output information (Pages 8-13).

Re claims 5 and 19, Shriberg teaches the method of claim 1, wherein determining the adjusted synthesized speech output (Pages 4-5, Why Use Prosody?) further comprises the steps of:

determining discourse function level prosodic feature adjustments (Pages 14-18);

However, Shriberg fails to teach determining input text;

determining the adjusted synthesized speech output based on the synthesized speech output and the discourse level prosodic feature adjustments

Lee teaches a TTS for interlocking with multimedia according to the present invention comprises a multimedia information input unit for organizing text, prosody, the information on synchronization with moving picture, lip-shape, and the information such as individual property; a data distributor by each media for distributing the information of the multimedia information input unit into the information by each media; a language processor for converting the text distributed by the data distributor by each media into

Art Unit: 2626

phoneme stream, presuming prosody information and symbolizing the information; a prosody processor for calculating a value of prosody control parameter from the symbolized prosody information using a rule and a table; a synchronization adjuster for adjusting the duration of the phoneme using the synchronization information distributed by the data distributor by each media; a signal processor for producing a synthesized speech using the prosody control parameter and data in a synthesis unit database; and a picture output apparatus for outputting the picture information distributed by the data distributor by each media onto a screen (Lee Col. 2 lines 29-49 & Fig. 1).

Lee further describes user intervention, wherein a user can change gender, age, and speech rate of the synthesized speech (Lee Col. 7 lines 10-17 & Fig. 1).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Shriberg in view of Marcu to incorporate determining input text and determining the adjusted synthesized speech output based on the synthesized speech output and the discourse level prosodic feature adjustments as taught by Lee to allow for the proper rules to analyze input text, wherein prosody control is established such as phonemic features of text in order to modify output speech synthesis based on input text, and words, sentences, or textual patterns in order to adapt in a changing environment (Lee Col. 7 lines 10-17 & Fig. 1).

Re claims 6 and 20, Shriberg teaches the method system of claim 1, wherein the model of discourse function level prosodic features (Pages 14-18) is a predictive model of discourse functions (Page 19).

Re claims 7 and 21, Shriberg teaches the method of claim 6, in which the predictive models are determined based on at least one of: machine learning and rules (Page 19).

Re claims 8 and 22, Shriberg teaches the method of claim 1, in which the prosodic features occur in at least one of a location: preceding, within and following the associated discourse function (Page 14).

Re claims 9 and 23, Shriberg teaches the method of claim 1, in which the prosodic features are encoded within a prosodic feature vector.

Re claims 10 and 24, Shriberg teaches the method of claim 9, in which the prosodic feature vector is a multimodal feature vector (Pages 14-18 & Table 10).

Re claims 11 and 25, Shriberg teaches the method of claim 1, in which the discourse functions include an intra-sentential discourse function (Page 8 & Table 1).

Re claims 12 and 26, Shriberg teaches the method of claim 1, in which the discourse functions include an inter-sentential discourse function (Page 8 & Table 1).

***Conclusion***

4. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael C. Colucci whose telephone number is (571)-270-1847. The examiner can normally be reached on 9:30 am - 6:00 pm, Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571)-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Art Unit: 2626

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Michael C Colucci/  
Examiner, Art Unit 2626  
Patent Examiner  
AU 2626  
(571)-270-1847  
[Michael.Colucci@uspto.gov](mailto:Michael.Colucci@uspto.gov)

/Richemond Dorvil/  
Supervisory Patent Examiner, Art Unit 2626